

AMENDMENTS TO THE CLAIMS

1. (currently amended) A processor-implemented method of collaborative focused crawling of documents related to focus topics on a network, the method comprising:

selectively prioritizing the documents to crawl based on a set of rules;

fetching prioritized documents from the network;

for each fetched document, determining whether the fetched document is relevant to any of the focus topics;

crawling the fetched document that matches any of the focus topics such that the fetched document is crawled only once even if the fetched document matches a plurality of the focus topics, wherein the fetched document comprises a document of interest for access by a user;

further crawling out-links on the fetched document based on an assumption that if the fetched document is of interest, the out-links are also of interest;

determining whether the fetched document should be disallowed, and upon determination that the fetched document should be disallowed:

selectively disallowing the fetched document;

identifying a resource locator string associated with the disallowed fetched document;

placing the resource locator string for the disallowed fetched document in a blacklist in order to prevent future crawling of the fetched document;

wherein the crawling is performed using a collaborative focus by analyzing the documents for more than one focus topic of the focus topics at a time by implementing a foreman function for reading a plurality of contents of the resulting uniform resource locator strings, and sequentially passing the contents of the resulting uniform resource locator strings to a plurality of sequentially configured miners that each represent an individual focus topic, wherein each miner is configured to instruct a fetcher to crawl a plurality of out-links on a document of the resulting resource locator string

whenever the contents of the resulting resource locator string match the specific focus topic of the miner, and wherein each miner is configured to ignore the resulting resource locator string whenever the contents of the resulting resource locator string do not match the specific focus of the miner; and

analyzing a plurality of fetched documents obtained from the crawling by providing a graphical indicia of one or more properties that are indicative of which of the plurality of fetched documents represent a best source of information for a specific topic within the plurality of fetched documents.

2. (original) The method of claim 1, further comprising seeding a plurality of seed uniform resource locator strings to start the collaborative focused crawling of the documents.

3. (original) The method of claim 2, further comprising crawling the seed uniform resource locator strings.

4. (original) The method of claim 3, further comprising writing a plurality of resulting uniform resource locator strings obtained by crawling the seed uniform resource locator strings.

5-9. (cancelled)

10. (currently amended) The method of claim 9, ~~further comprising~~
the wherein each miner allowing is individually configured to allow a crawling of the resulting resource locator string when the resulting resource locator string matches a plurality of web space rules.

11. (previously presented) The method of claim 10, wherein the web space

rules comprise domain rules, IP address rules, and prefix rules.

12. (currently amended) The method of claim 10, further comprising the plurality of miners collectively disallowing the crawling of the resulting resource locator string when the content of the resulting resource locator string ~~matches~~ fails to match a any of the focus topics of the miners.

13. (currently amended) The method of claim 10, wherein at least one of the miners comprises an unfocus miner that places the resulting uniform resource locator strings that match an unfocus topic in the blacklist, so that the uniform resource locator strings will not be crawled again.

14-20. (cancelled)